Appendix B: The National UIC Database

EPA Headquarters developed the National UIC Database, a well-centric database that collects and stores state[1] and regional direct implementation (DI) program data to support UIC programmatic data needs. The UIC system contains approximately 120 data elements related to various aspects of the UIC Program (e.g., permit information and enforcement and compliance data). The database includes a tool to electronically transfer data between existing state databases and the National UIC Database, a tool to generate standard reports, and a tool to support data quality assurance/quality control (QA/QC) reviews.

Headquarters deployed the National UIC Database in December 2007. As of July 2014, 14 programs have transitioned to e-reporting. For burden estimation purposes, 15 programs are expected to transition to e-reporting in 2015, 10 programs are expected to transition in 2016, and 11 programs are expected to transition in 2017.

To assist with funding challenges, EPA's Exchange Network (EN) Grants can be used to establish a data flow. Grant funds provide additional benefits to states for their own internal data management by allowing them to complete other related activities, such as building a state database or establishing an electronic data flow from owners or operators. Twenty-five UIC programs have already received funds to support activities to populate the National UIC Database. The EN Grant program provides funding, technical support, and tools to support these activities; programs have used funds to create a program database, improve the quality of source data, integrate state and federal data, and prepare internal systems to allow electronic data submission.

While the UIC regulations do not mandate the format in which the data must be submitted, historically, programs have provided information in paper format using the UIC program's 7520 reporting forms. States that transition to e-reporting are no longer required to submit the paper 7520 forms or complete web-based inventory and Program Activity Measures (PAMs) reporting; this eliminates the current burden associated with gathering and compiling data, completing the 7520 forms, and maintaining paper records. Although there is still burden associated with e-reporting, after establishing the data flow, the burden is less than that for paper based reporting. EPA developed tools to replicate the paper reports to ensure consistency in program management after transition to e-reporting. Programs have also reported data quality improvements as a result of the data flow process. The improvement in data quality may result in more efficient program management by both the states and EPA.

This Appendix describes the status of UIC data management by primacy states and DI programs and the burden and cost associated with data transfer.

B1. Burden and Cost Associated With Data Transfer

State UIC programs will incur burden and costs associated with the necessary start-up activities to prepare their databases and prepare data for transfer to Headquarters. Activities associated

---

[1] The term state includes states, tribes, and territories.

with data entry and semi-annual data transfers to Headquarters represent ongoing burden. EPA supports programs with funding through the EN grants, by providing technical assistance and training, and by creating individualized tools for programs to flow data.1

Burden associated with start-up activities

Prior to initiating data transfer, states must prepare UIC program applications and data to flow data to the National UIC Database. The effort and associated cost will vary depending on the status of the program's database.

Programs with Existing UIC Databases

Start-up costs include the costs associated with developing the data to transfer to the National UIC Database and setting up the data flow process. Activities include adding new data fields and new data tables (if needed) to align the state's UIC data to the national data model, cleaning up existing data, and in some cases, inputting historic data from records available in hard copy only.

On average, the total "data development" cost to these programs is about 310 hours. Annualizing the effort over an estimated 6-year phase-in schedule, this equals 51.6 hours per program, annually. Note that much of this is contractor labor, which is presented as non-labor costs in the UIC Programmatic ICR.

Programs must also set up the data flow infrastructure. This includes the following activities:

- Data mapping and data set generation, including mapping data from the state database structure to the National UIC Database structure and generating data sets with valid records,
- Convert to Extensible Markup Language (XML),[2] and
- Data flow via a National Environmental Exchange Network (NEIEN) Node involves setting up/testing automatic data submission and transfer to EPA's Central Data Exchange (CDX).

The total data flow burden is estimated to be 210 hours per program. Annualized over 6 years, this equates to 35 hours (in state and contractor labor) per year per program.

Programs with Little or no Electronic Data

States with little or no electronic data are assumed to require a less intensive effort than for those programs that already have a database as described above because there is less complexity in their data management needs. EPA estimates the start-up burden to these programs is 250 hours/ program (state and contractor labor), or 41.6 hours per year, annualized over 6 years.

Cost associated with start-up activities

---

[2] Extensible Markup Language (XML) is a set of rules for encoding documents in machine-readable form.

Start-up non-labor costs include contractor support and one-time hardware costs to set up a state NEIEN Node. EPA developed a Node Client that programs without a data node can use to submit UIC data. States may choose to develop an NEIEN Node to serve all state programs but it is no longer a cost they must incur to submit data to the National UIC Database.

EPA assumes that most start-up activities (estimated at 80 percent of the burden cited above), will be performed by contractors at an estimated labor rate of $91.58 hour. In addition, each UIC program will incur $7,000 ($1,167/year, annualized over 6 years) for costs associated with the development and maintenance of a state NEIEN Node to support the UIC data flow.

Ongoing activities

Annual activities include the incremental data entry burden for those programs that currently do not enter UIC data into a state database. (Programs that have a database are already incurring a data entry burden; thus, the ICR assumes no additional paperwork burden for these programs as a result of the new database.) The annual data entry burden is estimated to be 40 hours. Note that some programs have developed software that allow inspection data gathered in the field to be directly uploaded into their program database, further reducing data entry burden. Other programs have also established web-based permitting and inventory data entry applications that flow data to their program database further reducing burden.

All programs will submit their data to the National UIC Database through CDX and respond to QA and data validation issues. EPA estimates these tasks will require 4 hours per semi-annual submission, or 8 hours per year, per program.

B2. Burden Reduction and Cost Savings

Over the long run, the National UIC Database will reduce the programs' reporting burden. Following transfer to e-reporting, programs will no longer need to complete the 7520 forms and report on inventory and PAMs. This eliminates the current burden associated with gathering and compiling UIC data, completing the reporting forms, and maintaining records.

As described above, programs will need to make a time and hardware investment to develop the automated data transfer process. However, when the data transfer is complete, the current annual burden associated with completing the 7520 forms and reporting inventory and PAMs data will be replaced with a lower burden associated with placing data on the state NEIEN Node to flow it to Headquarters and providing QA to assure data accuracy, as described below:

- Under a paper-based system, programs would collectively incur over $625,000 annually in reporting and recordkeeping costs; this equals about $10,800 annually per state.

- Between 2015 and 2017, programs will collectively incur $1,497,209 in start-up costs to convert from paper-based to electronic reporting. This equates to an average of $49,382 in labor and hardware costs for each program that currently has a database, and $27,387 per program without a database.

- When their initial data transfer is complete, programs will incur lower annual costs: an estimated average annual cost over the ICR clearance period of about $80,860 for all programs collectively to enter data and transfer it to Headquarters semi-annually. Note that this unit burden is lower in this ICR, based on the results of a survey of e-reporting programs by Headquarters and a reduction in the frequency of data transfers.

Table B-1 presents the annual and cumulative costs to programs under paper based and electronic reporting systems. When the initial data population and transfer are complete, programs will collectively save over $500,000 annually.

Table B-1: Comparison of Costs – Paper-Based Reporting vs. Electronic Reporting				
	Paper-Based Reporting		**Electronic Reporting**	
Year	Annual cost	Cumulative cost	Annual cost*	Cumulative cost
2015	**$625,875**	**$3,755,251**	**$690,406**	**$1,458,480**
2016	**$625,875**	**$4,381,127**	**$485,374**	**$1,943,854**
2017	**$625,875**	**$5,007,002**	**$564,008**	**$2,507,862**
2018	$625,875	$5,632,877	$345,631	$2,853,494
2019	$625,875	$6,258,752	$115,324	$2,968,818
2020	$625,875	$6,884,627	$115,324	$3,084,143
2021	$625,875	$7,510,503	$115,324	$3,199,467
2022	$625,875	$8,136,378	$115,324	$3,314,791
2023	$625,875	$8,762,253	$115,324	$3,430,116
2024	$625,875	$9,388,128	$115,324	$3,545,440
2025	$625,875	$10,014,003	$115,324	$3,660,765
2026	$625,875	$10,639,879	$115,324	$3,776,089
* Includes the cost for new programs transitioning to e-reporting plus ongoing costs (e.g., for data entry and QA/QA) following transfer.				